Universal Code Equivalent of a Yeast Mitochondrial lntron Reading Frame Is Expressed into E. coli as a Specific Double Strand Endonuclease

Laurence Colleaux, L D'auriol, M Betermiert, G Cottarel, : Jacquier,', F. Galibert, B Dujorv

Universal Code Equivalent of a Yeast Mitochondrial Intron Reading Frame Is Expressed into E. coli as a Specific Double Strand Endonuclease

L. Colleaux,* L. d'Auriol,† M. Betermier,*‡
G. Cottarel,* A. Jacquier,*§ F. Galibert,† and B. Dujon*
* Centre de Génétique Moléculaire du CNRS
Laboratoire Associé à l'Université P. M. Curie
F-91190 Gif sur Yvette, France
† Laboratoire d'Hématologie Expérimentale
Centre Hayem
Hôpital Saint Louis
F-75010 Paris, France

Summary

The intron of the mitochondrial 21S rRNA gene of Saccharomyces cerevisiae (r1 intron) possesses a 235 codon long internal open reading frame (r1 ORF) whose translation product determines the duplicative transposition of that intron during crosses between intron-plus strains (omega⁺) and intron-minus ones (omega⁻). Using site-directed mutagenesis, we have constructed a universal code equivalent of the r1 ORF that, under appropriate promoter control, allows the overexpression in E. coli of a protein identical to the mitochondrial intron encoded "transposase". This protein exhibits a double strand endonuclease activity specific for the omega⁻ site. This finding demonstrates, for the first time, the enzymatic activity of an intron encoded protein whose function is to promote the spreading of that intron by generating double strand breaks at a specific sequence within a gene.

Introduction

Introns of mitochondrial and chloroplastic genes show two classes of potential RNA secondary structures believed to play an important role in splicing (Davies et al., 1982; Michel et al., 1982; Michel and Dujon, 1983; Waring and Davies, 1984). In yeasts and fungi, most mitochondrial introns also possess long internal open reading frames (ORFs), often linked with their preceding exons (reviewed in Dujon, 1983; Grivell, 1983). The translation products of such reading frames have not yet been biochemically characterized but their functions have been deduced from genetic data or inferred from DNA sequence comparisons, revealing a complex picture. First, messenger RNA "maturases" have been postulated for the products of some class I or class II intronic ORFs of Saccharomyces cerevisiae (Church and Gilbert, 1980; Jacq et al., 1980; Lazowska et al., 1980; Carignani et al., 1983). In such a model the intron encoded proteins are supposed to act at some essential step for proper RNA splicing to occur. Yet, the same class II intron ORF products reveal structural

features common to retroviral reverse transcriptases, suggesting their possible role in the propagation of introns (Michel and Lang, 1985). Second, a "transposase" function has been postulated for the ORF product of the class I intron of the 21S rRNA gene of S. cerevisiae (r1 intron) (Jacquier and Dujon, 1985; Macreadie et al., 1985). In such a model the intron encoded protein is supposed to be involved in the specific integration of that intron into a gene. Yet, the homologous intron in Neurospora crassa contains a different ORF which is believed to encode a mitoribosomal protein (Burke and RajBhandary, 1982). It is clear, at this stage, that a direct determination of the actual biochemical activities of such intron encoded proteins using in vitro assays is needed to ascertain these deductions.

We attempted to develop such an assay to study the biochemical function (or functions) of the r1 intron encoded transposase which has recently been shown to be responsible for the spreading of that intron into a gene (Jacquier and Dujon, 1985; Macreadie et al., 1985). This transposase is the product of a 235 codon long open reading frame (r1 ORF; Dujon, 1980) that is conserved in various Kluyveromyces species (Jacquier and Dujon, 1983). It is not required in the splicing of the 21S rRNA precursor molecules (Tabak et al., 1981), those being endowed with a self-splicing activity in in vitro tests (van der Horst and Tabak, 1985) similar to that of the class I intron of Tetrahymena rDNA (Kruger et al., 1982). The self spreading of the r1 intron results from a duplicative transposition, occurring in crosses between intron-plus strains (omega⁺) and intron-minus ones (omega⁻), that quantitatively converts intron-minus copies of the 21S rRNA gene into intron-plus copies. Mutants altering the transposition of the intron have been isolated at the recipient site where the r1 intron inserts into the omega⁻ 21S rRNA gene (omegaⁿ mutants; Dujon et al., 1976). This site (here called omega⁻ site) shows striking homology with the site of the HO-controlled double strand break in the yeast nuclear MAT gene (Strathern et al., 1982; Weiffenbach et al., 1983), suggesting the possibility that a double strand cut at the omega⁻ site initiates transposition of the r1 intron (Dujon and Jacquier, 1983), a fact that has recently been directly verified (Dujon et al., 1985; Zinn and Butow, 1985). Other mutants have been isolated within the r1 ORF itself (omegaᵈ mutants) demonstrating the direct role of this ORF product in the transposition and allowing speculation that this product recognizes the omega⁻ site and, possibly, generates itself the double strand cut at this site (Jacquier and Dujon, 1985; Macreadie et al., 1985).

An assay for the biochemical activity (or activities) of the r1 intron encoded transposase has, so far, remained elusive, for this protein is synthesized in such a low amount in mitochondria that it is not even detectable by standard analytical methods (Dujon et al., 1985), a feature common to other mitochondrial intron ORF products (Guiso et al., 1984; Jacq et al., 1984). To overcome this difficulty we decided to overexpress in E. coli a protein identical to the

Table 1. Codon Usage in the r1 ORF and Its Universal Code Equivalent

	r1 M	r1 U.C.	Codon	Expected Number of Codons			r1 M	r1 U.C.	Codon	Expected Number of Codons	
F	11	0	UUU / UUC	2.4 / 8.6	F	Y	14	1	UAU / UAC	3.3 / 11.7	Y
L	26 → 22	0	*UUA* / UUG	0.2 / 0.5		H	4	0	CAU / CAC	0.7 / 3.3	H
T	2 / 0		CUU / CUC	1.0 / 1.3	L	Q	10	0	CAA / CAG	1.2 / 8.8	Q
	1 → 0	0 → 4	*CUA* / CUG	0 / 22.6		N	24	1	*AAU* / AAC	1 / 24	N
	4 → 5 / 1 / 2 / 0		ACU / ACC / ACA / ACG	3.7 / 5.2 / 0.3 / 0.7	T	K	25	1	AAA / AAG	19.2 / 6.8	K
I	18	0	AUU / AUC	3.1 / 14.9	I	D	8	0	GAU / GAC	2.7 / 5.3	D
M	5 → 0	4 → 9	*AUA* / AUG	0 / 9	M	E	11	1	GAA / GAG	9.2 / 2.8	E
V	4 / 0 / 7 / 0		GUU / GUC / GUA / GUG	5.5 / 0.8 / 3.0 / 1.8	V	C	3	0	UGU / UGC	1 / 2	C
S	3 / 1 / 3 / 0 / 5 / 0		UCU / UCC / UCA / UCG / AGU / AGC	4.3 / 4.4 / 0.2 / 0.5 / 0.1 / 2.4	S	W	5 → 0 / 0 → 5		*UGA* / UGG	0 / 5	Stop / W
P	8 / 0 / 1 / 0		CCU / CCC / CCA / CCG	0.5 / 0 / 1.2 / 7.3	P	R	2 / 0 / 0 / 0 / 2 / 0		CGU / CGC / *CGA* / *CGG* / *AGA* / *AGG*	2.9 / 1.1 / 0 / 0 / 0 / 0	R
A	4 / 0 / 3 / 0		GCU / GCC / GCA / GCG	2.7 / 0.6 / 1.9 / 1.9	A	G	6 → 7 / 0 / 4 → 3 / 0		GGU / GGC / *GGA* / GGG	5.7 / 4.1 / 0 / 0.2	G

This table indicates for each codon: the corresponding amino acid in the mitochondrial genetic code of S. cerevisiae (left) or in the universal code (right); the number of occurrences of that codon in the mitochondrial wild-type r1 ORF (M) or in its universal code equivalent constructed in this work (U.C.). The expected occurrences of each codon have been calculated for a theoretical gene encoding the same polypeptide, using the codon bias known for E. coli major proteins (Gouy and Gautier, 1982; Gouy, personal communication). Codons very rarely used in major E. coli proteins are italicized.

expected genuine mitochondrial protein. Because the yeast mitochondrial genetic code differs from the universal code, we constructed a universal code equivalent of the r1 ORF of S. cerevisiae, using oligonucleotide-directed mutagenesis. This universal code equivalent efficiently directs the synthesis, in E. coli, of a full length protein which bears all characteristics predicted for the native mitochondrial protein. We show that this protein generates a specific double strand cut at the omega⁻ site.

Results

Construction of a Universal Code Equivalent from the Mitochondrial r1 ORF

The genetic code in mitochondria differs from the universal code (Barrel et al., 1979). In the case of S. cerevisiae

the differences are: the use of the stop codon TGA as a tryptophane codon (Fox, 1979; Macino et al., 1979); the use of the isoleucine codon ATA to specify methionine (Hudspeth et al., 1982); and the use of the leucine codon family CTN to specify threonine (Li and Tzagoloff, 1979; Bonitz et al., 1980). The r1 ORF contains 13 nonuniversal codons (five TGA, five ATA, two CTT, and one CTA) as well as a number of codons (approximately 30% of all codons) that are very rarely found in E. coli genes encoding major proteins (see Table 1). In order to obtain an efficient synthesis in E. coli of a polypeptide identical in size and in sequence to the intron encoded transposase expected to be synthesized in yeast mitochondria, we constructed a universal code equivalent of the r1 ORF by oligonucleotide-directed mutagenesis. To do this, nine different oligonucleotides were synthesized (Table 2) to change the five

Table 2. The Synthetic D Oligonucleotides Used for In Vitro Mutagenesis

		Dissociation Temperatures						Nucleotides Changed and Mismatches	Codons Changed and Positions
		Calculated				Experimental			
		a	b	c	d	e	f		
D1	3' TAAACTCACCTTTTTATTC 5'	43	49	44	48	45	60	1(C/A)	TGA → TGG 62
D2	3' ACTAGTTACCCATAATAGT 5'	45	51	46	50	45	60	1(C/A)	TGA → TGG 78
D3	3' ATCGTATAACCAAATACCTACTACCTCCATTTACCCTAATATTA 5'	67	75	ND	ND	66	79	3(C/A)	TGA → TGG 141 ATA → ATG 143 TGA → TGG 149
D4	3' CATTAATGGACCCCTCGAGTT 5'	55	61	58	62	57	66	1(C/A)	TGA → TGG 99
D5	3' GTTTTTTATTTTAGTATACTTTTTATAA 5'	48	56	56	62	57	66	2(G/A; A/A)	NdeI site 0
D6	3' GAGTTTACTACATATTTGACGGATTATGAT 5'	54	66	66	78	54	70	3(C/A; G/T)	ATA → ATG 221 TTA → CTG 224
D7	3' AACTAAGATACTCAATAGACTAAAAAATAT 5'	50	62	60	72	52	69	3(C/A; G/T)	ATA → ATG 203 TTA → CTG 206
D8	3' ATTGTGGATACTCAGACCGTATAACCAAAT 5'	(w) 53 (m) 57	69 69	66 70	82 82	51 55	 72	4(C/A; G/T) 3(C/A; G/T)	ATA → ATG 136 TTA → CTG 138
D9	3' TAGTTCATTACTTAGACCCAGGATTAAGAT 5'	55	67	66	80	51	72	4(C/A; G/T; A/A)	ATA → ATG 10 TTA → CTG 12 GGA → GGT 13
D10	3' ACTTCCATTTTGAATAACATACG 5'	42	58	52	60	54	67	3(C/T; G/T; A/A)	CTA → ACT 55
						TOTAL		24	16

The sequences of the 10 oligonucleotides are given in their 3' to 5' orientation to facilitate alignment with the r1 ORF sequence (see Figure 2). Each oligonucleotide has been synthesized complementary to the mRNA of the wild-type r1 ORF with the exception of the mutation (or mutations) introduced (underlined characters). Dissociation temperatures (°C) between each oligonucleotide and either the wild-type r1 ORF (a, c, and e) or its universal code equivalent (b, d, and f) were calculated according to Lathe (1985) (a and b) or to Suggs et al. (1981) (c and d) or were determined experimentally (e and f) as described in Experimental Procedures. All dissociations temperatures refer to the conditions used (6 × NET buffer). For the D8 oligonucleotide, which overlaps the D3 one, dissociation temperatures are given against the wild-type sequence (w) as well as against the D3 mutant sequence (m).
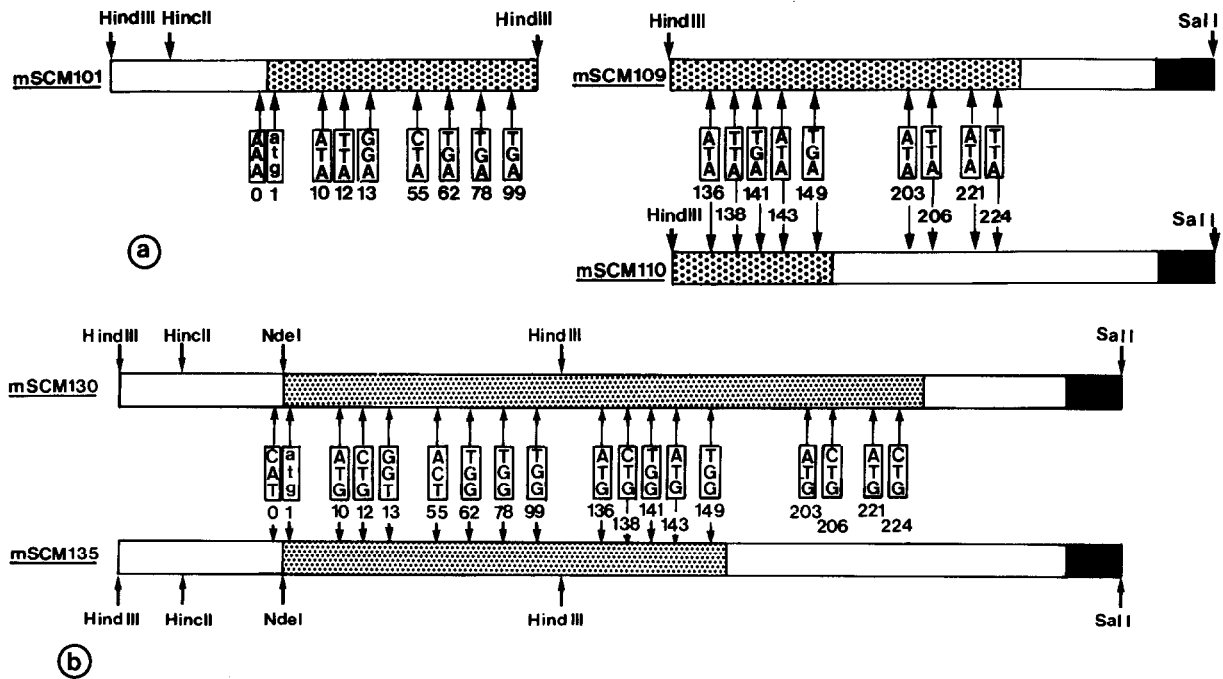
Figure 1. Strategy for the Construction of the Universal Code Equivalent of the r1 ORF

Full sectors, downstream exon of the 21S rRNA gene; dotted sectors, the r1 ORF; void sectors, segments of the r1 intron flanking the r1 ORF. All modified codons are indicated with their position. The ATG initiator codon, although not changed, is indicated for memory. The CAT change introduced at position 0 refer to the Nde I site created.

(a) Fragments of the r1 ORF introduced in M13 vectors as starting material for the stepwise oligonucleotide directed mutagenesis (see Table 3). The upstream part of the r1 ORF is carried on a 495 bp long Hind III–Hind III fragment introduced into the unique Hind III site of the vector M13mp8 to produce the recombinant mSCM101 (this fragment originates from yeast strain KM227-1A; see Jacquier and Dujon, 1985). The downstream part of the r1 ORF is carried on a 581 bp long Hind III–Sal I fragment, introduced between the Hind III and the Sal I sites of the vector M13mp9. This part contains the −1 frameshift in the omega^d mutant and has, therefore, been mutagenized in parallel on the recombinant mSCM109 (carrying the omega^+ fragment originating from the yeast strain IL8-8C; see Dujon, 1980) and on the recombinant mSCM110 (carrying the omega^d fragment originating from the yeast strain KM227-1A; see Jacquier and Dujon, 1985).

(b) Reconstruction of the complete r1 ORFs from the M13 recombinants resulting from the oligonucleotide directed mutagenesis (see Table 3). The Hind III–Hind III fragment was purified from the replicative form of mSCM117 and ligated into the unique Hind III site of the recombinant phages mSCM124 and mSCM126 to reconstruct the complete r1 ORFs, in their universal code version, from the omega^+ wild type (mSCM130) or from the omega^d recombinant (mSCM135).

TGA codons into TGG codons, the five ATA codons into ATG codons, and the unique CTA codon into an ACT codon (a major threonine codon in E. coli). Only this codon of the CTN family has been changed because this threonine was known to be conserved in Kluyveromyces thermotolerans (and coded for by ACT) while the two others are replaced, in this last species, by CAT (histidine) and AAT (asparagine), respectively (Jacquier and Dujon, 1983). To better adapt the universal code equivalent constructed to the E. coli translation machinery, some of the oligonucleotides were synthesized such as to simultaneously replace some of the very rare codons by synonymous codons more common to E. coli. In addition, one oligonucleotide was synthesized to introduce a unique restriction site (Nde I) at the natural ATG initiator codon of the r1 ORF to subsequently permit expression from this codon after appropriate ligations of the r1 ORF universal code equivalent into E. coli expression vectors.

To provide a specific control for the final biochemical assays of the E. coli made protein, two constructs were planned in parallel, one from the r1 ORF of a wild-type omega^+ yeast strain and the other from the r1 ORF of an omega^d frameshift mutant (Jacquier and Dujon, 1985). In order to optimize the in vitro mutagenesis, the entire r1 ORF was first introduced in two parts in M13 vectors, the upstream part which is common to both the omega^+ and the omega^d r1 ORF being mutagenized only once (Figure 1a). After the successive cycles of oligonucleotide-directed in vitro mutagenesis were performed on each fragment (Table 3), the complete universal code equivalents of the omega^+ r1 ORF and of the omega^d r1 ORF were reconstructed (Figure 1b). The DNA sequences of the entire 1076 bp long Hind III–Sal I fragments of the two recombinants isolated (mSCM130 and mSCM135) were ultimately verified by the chain termination method using the D oligonucleotides as successive primers to ensure continuous reading of overlapping segments. In total, mSCM130 and mSCM135 harbor respectively 24 and 18 nucleotide substitutions from the original mitochondrial sequence (corresponding to the replacement of 16 and 12 codons respectively). Figure 2 gives the final sequence of the universal code equivalent of the r1 ORF used in subsequent experiments and the corresponding amino acid sequence.

Table 3. Stepwise Oligonucleotide-Directed Series of Mutagenesis

Cycle Number	Oligonucleotide	Recombinant Clone	Codon Number									Total Nucleotides Changed	Total Codons Changed
			0	1	10	12	13	55	62	78	99		
		mSCM 101	AAA	ATG	ATA	TTA	GGA	CTA	TGA	TGA	TGA	0	0
1	D1 + D2	mSCM 106							TGG	TGG		2	2
2	D5	mSCM 108	CAT						TGG	TGG		4	2
3	D4	mSCM 113	CAT						TGG	TGG	TGG	5	3
4	D10	mSCM 116	CAT					ACT	TGG	TGG	TGG	8	4
5	D9	mSCM 117	CAT		ATG	CTG	GGT	ACT	TGG	TGG	TGG	12	7

Cycle Number	Oligonucleotide	Recombinant Clone	Codon Number									Total Nucleotides Changed	Total Codons Changed
			136	138	141	143	149	203	206	221	224		
		mSCM 109	ATA	TTA	TGA	ATA	TGA	ATA	TTA	ATA	TTA	0	0
1	D6	mSCM 121								ATG	CTG	3	2
2	D7	mSCM 122						ATG	CTG	ATG	CTG	6	4
3	D3	mSCM 123			TGG	ATG	TGG	ATG	CTG	ATG	CTG	9	7
4	D8	mSCM 124	ATG	CTG	TGG	ATG	TGG	ATG	CTG	ATG	CTG	12	9

Cycle Number	Oligonucleotide	Recombinant Clone	Codon Number					Total Nucleotides Changed	Total Codons Changed
			136	138	141	143	149		
		mSCM 110	ATA	TTA	TGA	ATA	TGA	0	0
1	D3	mSCM 125			TGG	ATG	TGG	3	3
2	D8	mSCM 126	ATG	CTG	TGG	ATG	TGG	6	5

The M13 recombinant clone at the origin of each series is underlined along with the corresponding codons and their positions. The table indicates all codons of the r1 ORF that have been mutagenized after each cycle (using the indicated D oligonucleotide) to produce the mutant clone indicated (mSCM number). Positions of codons and cumulative numbers of nucleotides and codons changed in each mutant clone are indicated respectively to the original r1 ORF sequence.

Ndel

```
Ndel            10              20              30
    M K N I K K N Q V M N L G P N S K L L K E Y K S Q L I E L N I E Q F E A G I G
CATATGAAAATATTAAAAAAATCAAGTAATGAATCTGGGTCCTAATTCTAAATTATTAAAAGAATATAAATCACAATTAATTGAATTAAATATTGAACAATTTGAAGCAGGTATTGGT
AA                          A  T A A

40              50              60              70
    L I L G D A Y I R S R D E G K T Y C M Q F E W K N K A Y M D H V C L L Y D Q W V
TTAATTTTAGGAGATGCTTATATTCGTAGTCGTGATGAAGGTAAAACTTATTGTATGCAATTTGAGTGGAAAAATAAGGCATACATGGATCATGTATGTTTATTATATGATCAATGGGTA
                              CTA                   A                                              A

80              90              100             110
    L S P P H K K E R V N H L G N L V I T W G A Q T F K H Q A F N K L A N L F I V N
TTATCACCTCCTCATAAAAAAGAAAGAGTTAATCATTTAGGTAATTTAGTAATTACCTGGGGAGCTCAAACTTTTAAACATCAAGCTTTTAATAAATTAGCTAACTTATTTATTGTAAAT
                                                                              A

120             130             140             150
    N K K(L)I P N N L V E N Y L T P M S L A Y W F M D D G G K W D Y N K N S(L)N K S
AATAAAAAACTTATTCCTAATAATTTAGTTGAAAATTATTTAACACCTATGAGTCTGGCATATTGGTTTATGGATGATGGAGGTAAATGGGATTATAATAAAAATTCTCTTAATAAAAGT
                                      A  ↓A      A      A              A                          ⇧

160             170             180             190
    I V L N T Q S F T F E E V E Y L V K G L R N K F Q L N C Y V K I N K N K P I I Y
ATTGTATTAAATACACAAAGTTTTACTTTTGAAGAAGTAGAATATTTAGTTAAAGGTTTAAGAAATAAATTTCAATTAAATTGTTATGTTAAAATTAATAAAAATAAACCAATTATTTAT

200             210             220             230
    I D S M S Y L I F Y N L I K P Y L I P Q M M Y K L P N T I S S E T F L K ***
ATTGATTCTATGAGTTATCTGATTTTTTATAATTTAATTAAACCTTATTTAATTCCTCAAATGATGTATAAACTGCCTAATACTATTTCATCCGAAACTTTTTTAAAATAA
          A    T A                                              A    T A
```
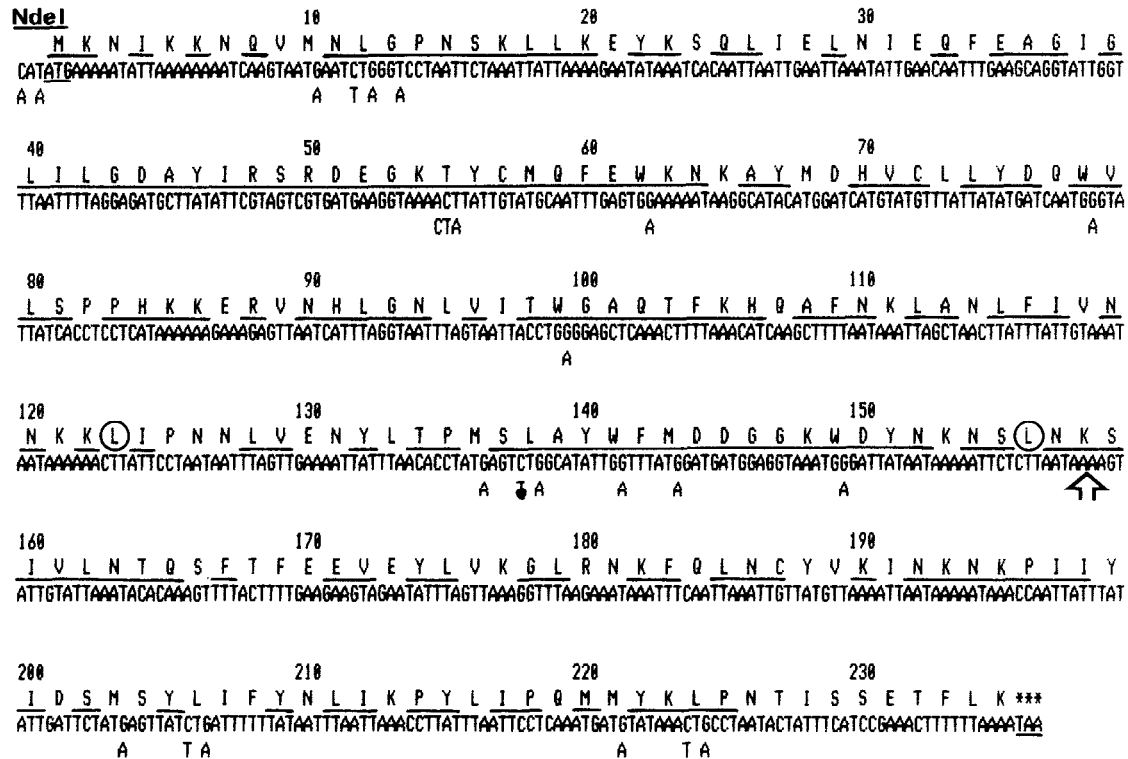
Figure 2. Sequence of the Universal Code Equivalent of the r1 ORF

The sequence of the DNA strand identical to the mRNA is given from the unique Nde I site, created at the initiator ATG codon, to the terminal TAA codon. Differences in the original mitochondrial sequence are indicated below the universal code equivalent. The corresponding amino acid sequence, translated according to the universal code, is identical to that of the mitochondrial r1 ORF translation product translated according to the mitochondrial code except for the two threonines at positions 123 and 156 which are replaced by two leucines (circles). Underlined amino acids are conserved in Kluyveromyces thermotolerans (Jacquier and Dujon, 1983). The −1 frameshift mutation of the omega^d mutant (arrow) occurs at codon 158 (Jacquier and Dujon, 1985). The universal code equivalent of the omega^d r1 ORF is identical to that of the omega^+ r1 ORF upstream of that codon and identical to the original mitochondrial sequence downstream of that codon.
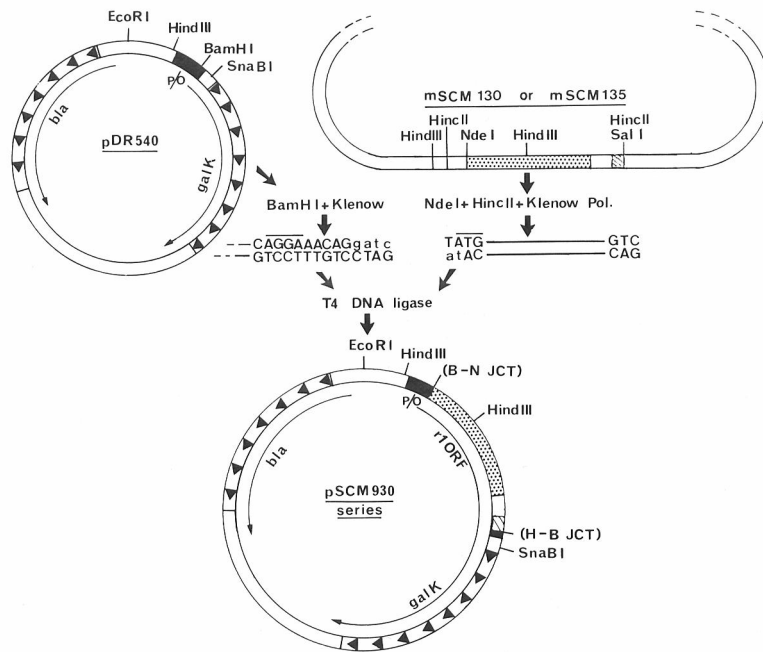
Figure 3. Construction of the Expression Plasmids of the pSCM930 Series

Top: map of the vector pDR540 and of the M13 clones carrying the universal code equivalent of the r1 ORF (dotted sector). The *tacI* promoter and the adjacent ribosome binding site are indicated by the full sector (P/O). Extent and orientation of the beta-lactamase gene (bla) and of the GalK gene are indicated. Hatched sectors: the downstream exon of the 21S rRNA gene.

Middle: the 904 bp long Nde I–Hinc II fragment from mSCM130 and the corresponding 903 bp long fragment from mSCM135 were purified and the Nde I staggered cuts were converted to blunt ends using the Klenow polymerase. Blunt end ligations of these fragments into the Klenow-filled Bam HI site of pDR540 places, in the proper orientation of the insert, the natural ATG initiator codon of the r1 ORF at an appropriate distance from the Shine-Dalgarno sequence of the vector.

Bottom: map of the resulting series of plasmids (plasmids number pSCM930, 933, 934, and 935 contain the *omega+* r1 ORF universal code equivalent; plasmids pSCM931 and pSCM936 contain the *omega^d* r1 ORF universal code equivalent).

Expression in E. coli of the Universal Code Equivalent of the r1 ORF

The complete universal code equivalents from both the *omega+* wild-type and the *omega^d* mutant of the r1 ORF have been introduced into the E. coli expression vector pDR540 under the control of the *tacI* promoter in order to allow expression of the protein from its natural ATG initiator codon, using the ribosomal binding site of the vector (Figure 3). Notice that this construct creates an artificial operon under the control of the *tac* promoter by placing the r1 ORF equivalent in front of the GalK gene so that comparison of the levels of expression of the two translation products in E. coli should directly reveal posttranscriptional steps.

Expressions of the r1–ORF universal code equivalent in the recombinant plasmids of the pSCM930 series were first assayed in IPTG-induced E. coli maxicells as described in Experimental Procedures. Figure 4 shows the results for four independent recombinant plasmids containing the *omega+* r1 ORF equivalent (pSCM930, 933, 934, and 935) and two independent plasmids containing the *omega^d* r1 ORF equivalent (pSCM931 and 936). The most prominent band (42 kd), which is present in all recombinants as well as in the vector alone, corresponds to the galK protein. The second band by labeling intensity corresponds exactly to the electrophoretic mobility expected for the full length r1 ORF translation product, initiated at the natural ATG codon (235 amino acids long). Identification of this 25.5 kd protein as the r1 ORF product is further confirmed by its disappearance in the recombinant plasmids carrying the *omega^d* r1 ORF and its replacement by a 17.7 kd protein which, again, exactly corresponds to the size expected for a translation product
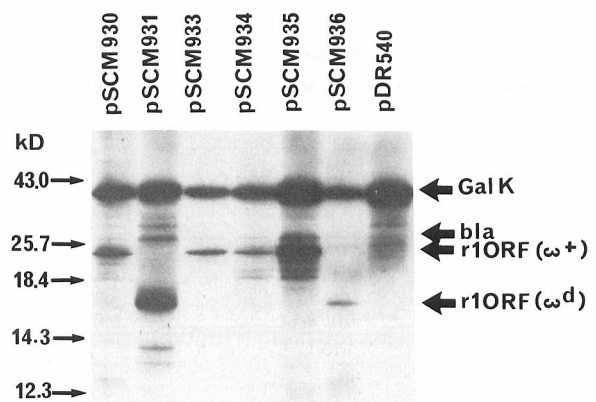

Figure 4. Expression of the pSCM930 Series in E. coli Maxicells

Extracts of ^{35}S-methionine-labeled E. coli maxicells containing various plasmids were analyzed on a 15% SDS urea PAG. The gel was calibrated using prestained "low molecular weight protein standards" from B. R. L. (kd: kilodaltons). After fixation the dried down gel was autoradiographed on Kodak XAR-5 film using EN³HANCE amplifier (New England Nuclear). Major ^{35}S-labeled translation products were identified from their electrophoretic mobilities as the product of the galK gene, the beta-lactamase gene (bla) and the r1 ORF universal code equivalents from *omega+* or from *omega^d*.

starting at the same natural ATG initiator codon and ending at the TAA codon created by the −1 frameshift mutation (161 amino acids long). The ^{35}S label intensities of the r1 ORF translation products were compared to the GalK translation product by scintillation counting of the bands cut off the gels (both proteins contain nine methionines). Although upstream on the polycistronic transcript, the r1 ORF product amounts to only 20%–50% of the
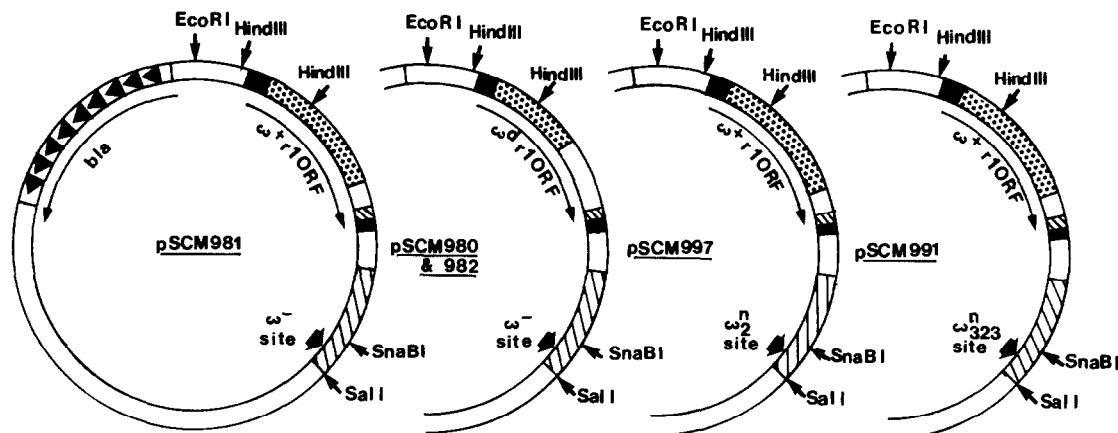
Figure 5. Construction of the Self Linearizable Plasmids of the pSCM980 and pSCM990 Series

The pSCM980 series: the 1371 bp Eco RI–Sna BI fragment from pSCM930 and the 4598 bp Eco RI–Eco RV fragment from pSCM529 (Dujon, 1980) were gel purified and ligated. One recombinant plasmid (pSCM981) that contains the universal code equivalent of the $omega^+$ r1 ORF was picked up (the unique sticky end at the Eco RI site determines the orientation of the insert). The same construction was performed in parallel using the 1370 bp Eco RI–Sna BI fragment from plasmid pSCM931 to yield the recombinant plasmids pSCM980 and pSCM982 that contain the universal code equivalent of the $omega^d$ r1 ORF. Same symbols as Figure 3; hatched sectors, exons of the 21S rRNA gene.

The pSCM990 series: the 132 bp long Sna BI–Sal I fragment of pSCM981 was replaced by the same fragment issued from the mtDNA of the $omega^n$ strain IL324-3A/2 (Dujon and Jacquier, 1983) or from the plasmid pSCM511 (Jacquier and Dujon, 1983). One recombinant plasmid (pSCM997) containing the $omega^n_2$ mutant site and one recombinant plasmid (pSCM991) containing the $omega^n_{323}$ site were picked up. The unique sticky end at the Sal I site determines the orientation of the insert.

GalK gene product in the various experiments performed. The beta-lactamase gene, which is carried on the same plasmid (see Figure 3), but under the control of another promoter, is expressed to a much lower level (the protein contains 10 methionines). Additional bands of lower intensities are also visible on Figure 4 (especially in the lanes that contain the highest total radioactivity). Some of them are also found in the vector and may correspond to the background of E. coli proteins. The others, ranging in size from 21 to 24 kd, reproducibly appear in the plasmids containing the $omega^+$ r1 ORF only. They are also visible on the nonoverexposed lanes and may represent premature chain terminations of the r1 ORF or specific degradation products of this protein.

We conclude from these experiments that our universal code equivalent of the r1 ORF is efficiently expressed into E. coli as a polypeptide that corresponds in size to the one expected to be made in yeast mitochondria. This polypeptide is sufficiently stable to be present in a relatively important amount in the cell. Its sequence should be identical to the one in mitochondria except for the two amino acids coded for by the two CTT codons that have not been changed (see Figure 2). The truncated polypeptide synthesized by the $omega^d$ r1 ORF is similarly stable and also corresponds in size to the one expected to be made in mitochondria of the $omega^d$ mutant.

The r1 ORF Protein Synthesized in E. coli Shows a Double Strand Endonuclease Activity Specific for the $omega^-$ Site

Because our previous genetic evidence (Jacquier and Dujon, 1985; Dujon et al., 1985) indicated that the r1 ORF translation product is required, in mitochondria, to generate the double strand cut observed at the $omega^-$ site,

we have first assayed the E. coli synthesized r1 ORF protein for its possible endonuclease activity on the $omega^-$ site. To do this, we have constructed two new series of plasmids (Figure 5) that permit a simple in vivo self linearization assay. The first series of plasmids (pSCM980 series) contains simultaneously the normal $omega^-$ site and the universal code equivalent of either the $omega^+$ r1 ORF (pSCM981) or the $omega^d$ r1 ORF (pSCM980 and pSCM982). The second series of plasmids (pSCM990 series) contains simultaneously the universal code equivalent of the $omega^+$ r1 ORF and either the $omega^n_2$ mutant site (pSCM997) or the $omega^n_{323}$ mutant site (pSCM991).

All plasmids of the pSCM980 and pSCM990 series were tested for the expression of their r1 ORFs in E. coli maxicells (data not shown). We found that they determine the efficient synthesis of the expected r1 ORF product with no significant differences in the levels of expression as compared to the plasmids of the pSCM930 series. The in vivo self linearization assays were then performed with the above plasmids in the E. coli strain DH1 by testing the occurrence of a double strand cut at the $omega^-$ site upon induction of the tac promoter. Figure 6a shows the rationale of the experiments.

In a first series of experiments, plasmids were extracted using the "gentle" procedure from midlogarithmic cultures of transformed E. coli cells grown either in the absence or in the presence of IPTG (see Experimental Procedures). DNA preparations were then restricted using a variety of endonucleases (Figure 6a), electrophoresed, and hybridized with the mitochondrial 21S rRNA gene as a probe. Figures 6b and 6c show examples of such a test using the two restriction endonucleases Hinf I and Hind III. It can be seen that, in addition to the expected fragments gener-
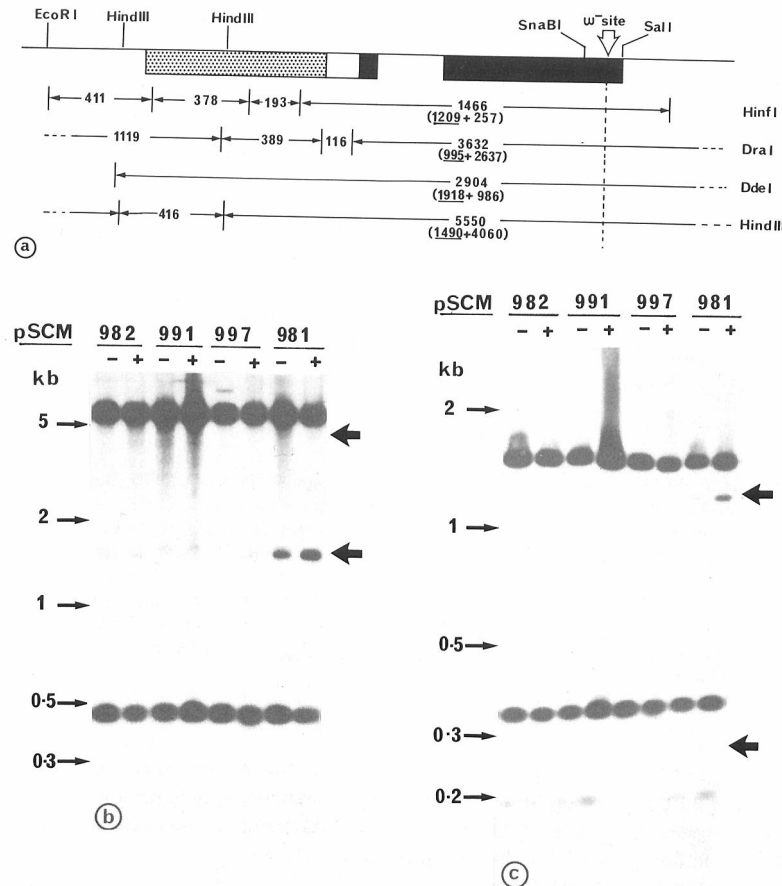
Figure 6. In Vivo Assay of the Endonuclease Activity of the r1 ORF Translation Product

(a) Partial restriction map of the plasmids of the pSCM980 and pSCM990 series showing the fragments that hybridize with the mitochondrial 21S rRNA gene used as a probe. Full sectors, exons of the 21S rRNA gene; dotted sector, the r1 ORF; void sector, the downstream part of the r1 intron. Figures represent length (in base pairs) of restriction fragments generated by each restriction endonuclease used. Figures in parentheses indicate the new restriction fragments generated if a double strand cut is generated at the omega⁻ site (notice that one of the two new fragment share only 53 nucleotides of homology with the probe while the other [underlined] has a long segment of homology, hence their difference in intensity of hybridization).

(b and c) E. coli cells (DH1 strain) containing the plasmids indicated by the pSCM numbers were grown with (+) or without (−) IPTG as described in Experimental Procedures. Restriction digests of total DNA preparations made by the gentle procedure (see Experimental Procedures) from such cells were electrophoresed on a 1% agarose gel (Hind III digests: b) or a 1.5% agarose gel (Hinf I digests: c), transferred to nitrocellulose, and hybridized using ³²P-labeled purified mtDNA of the strain HC9-7/J262 as a probe (this strain is a rho⁻ derivative containing the entire type 2 21S rRNA gene; see Jacquier and Dujon, 1985). The arrows point to the positions of the additional fragments generated by the double strand cut at the omega⁻ site. kb: kilobase pairs.

ated by each restriction enzyme, new fragments appear for the plasmid pSCM981. These additional fragments exactly correspond to the sizes expected if a double strand cut is generated at the omega⁻ site. A similar situation was also found for the two other restriction endonucleases tested (data not shown). Additional fragments are generally less intense in the noninduced E. coli cells than in the induced ones, but they are nevertheless present. This is because the tac promoter is not entirely repressed in the absence of IPTG in the E. coli strain used. No similar additional fragments can be detected with the plasmid pSCM982 or with the plasmids pSCM991 and pSCM997. We conclude, therefore, that a double strand cut is generated in vivo at the omega⁻ site upon induction of the expression of the omega⁺ r1 ORF universal code equivalent. When the omega⁻ site is replaced by either omegaⁿ site (plasmids pSCM991 and pSCM997), no double strand cut occurs upon induction of the same r1 ORF. Thus, the specificity of the double strand endonuclease activity observed in E. coli is identical to that predicted for the r1 ORF product on the basis of genetic results in yeast mitochondria (Jacquier and Dujon, 1985). When the omega⁺ r1 ORF is replaced by the omegaᵈ r1 ORF (plasmid pSCM982) no double strand endonuclease activity can be detected on the normal omega⁻ site upon IPTG induction (the very low amount of linearized plasmids that seem to

be present in the Hind III digest [Figure 6b] most probably represents some readthrough of the omegaᵈ frameshift mutation by E. coli and not a residual activity of the truncated protein). This result is, again, in complete agreement with the genetic results in yeast mitochondria and demonstrates, therefore, that the specific double strand endonuclease activity observed in E. coli does result from the normal r1 ORF product synthesized in E. coli.

In a second series of experiments, we examined the kinetics of induction of the endonuclease activity of the r1 ORF product in parallel cultures of E. coli cells containing the plasmids pSCM981 or pSCM980. Figure 7a shows the growth curves of the E. coli cultures in the various conditions used. It can be seen that no significant difference exists between induced or noninduced conditions and between the two plasmids. Samples taken at early logarithmic phase (1), mid-logarithmic phase (2), or late logarithmic phase (3) were analyzed as described above. Figure 7b shows that, in the case of pSCM981, plasmid DNA molecules having a double strand cut at the omega⁻ site are already present at early logarithmic phase. The proportion of such in vivo self-linearized plasmids is higher for the IPTG induced culture than for the noninduced ones, as can be seen from the intensity of the additional bands relative to that of normal bands. This proportion increases during subsequent incubation of the cells
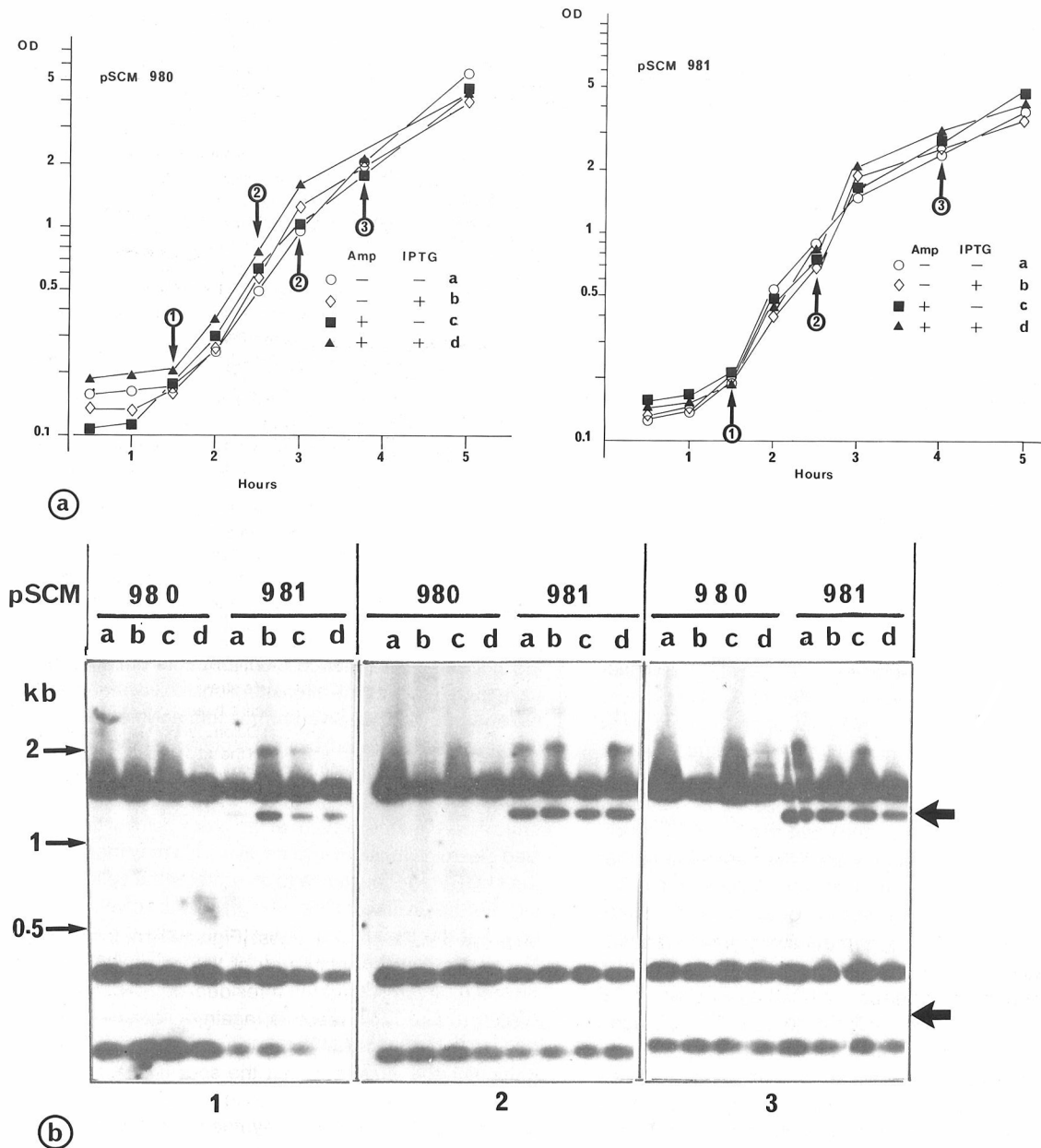
Figure 7. Kinetics of Induction of the r1 ORF Endonuclease Activity in E. coli

One milliliter aliquots of freshly grown E. coli cells (DH1 strain) containing the plasmid pSCM980 or the plasmid pSCM981 were inoculated in parallel into 100 ml LB medium with (+) or without (−) ampicillin (25 μg/ml) and IPTG (24 mg/ml) (a,b,c, and d conditions) and cultures were incubated at 37°C with aeration. Optical density of cultures (650 nm) were measured after appropriate dilutions and samples of appropriate volumes to compensate for cell density were withdrawn at intervals (early log, 1; middle log, 2; and late log phase, 3). Total DNA preparations made by the gentle method, were digested by Hinf I, electrophoresed on a 1.5% agarose gel, and hybridized using ^{32}P-labeled mtDNA of strain HC9-7/J262 as probe (see Figure 6a).

up to the point where no significant difference exists between induced and noninduced cultures. However, the self-linearization of the plasmid never goes to completion even after prolonged incubation of the induced cultures in stationary phase (data not shown). Figure 7b also shows that, when similar samples are taken from the cultures containing the plasmid pSCM980, no double strand cut at the omega⁻ site can be detected in any conditions. This confirms that the omega⁽ᵈ⁾ r1 ORF remains inactive in all

conditions assayed in which the omega⁺ r1 ORF is active, hence confirming that the double strand endonuclease activity can be attributed to the r1 ORF translation product. The kinetic experiment described also shows that the activity of the r1 ORF translation product has no measurable deleterious effect on the E. coli cells even in the conditions in which an appreciable fraction of plasmids are linearized. We have, finally, verified that the double strand cut at the omega⁻ site occurs in vivo prior to

plasmid purification (and not during purification steps or even during restriction endonuclease incubation of the DNA preparations) by including a treatment by proteinase K and phenol in the gentle preparation procedure immediately after cell lysis. Such a treatment shows no significant difference in the relative proportion of additional fragments recovered. In addition, when the fraction of supercoiled plasmids is purified from IPTG induced cells containing pSCM981, no double strand cut is generated at the omega⁻ site after analysis as described previously.

Discussion

The Universal Code Equivalent of the r1 Intron Reading Frame Is Efficiently Expressed in a Mature Size Protein in E. coli

The aim of our work was the efficient synthesis, in E. coli, of a protein identical to the expected mitochondrial r1 ORF translation product in order to examine its biochemical activity. Because the genuine mitochondrial protein has so far remained undetectable, arguments that the E. coli made protein faithfully represents the mitochondrial protein can only be deductive. In our case, however, the previous genetic and molecular characterization of the r1 intron help to solve this problem. First, the exact N-terminal amino acid of the genuine mitochondrial r1 ORF product can be predicted with no ambiguity since the only open reading frame of the r1 intron is entirely included within the intron itself (a rare feature for mitochondrial intron ORFs that are generally linked in phase with their preceeding exon; see Dujon, 1981) and the first ATG codon that occurs three triplets only after the beginning of that ORF is conserved in K. thermotolerans (Jacquier and Dujon, 1983). Second, the existence of an omega^d mutant provides us with a shorter and inactive protein as a control. Because the sizes of the E. coli translation products exactly correspond (within the limits of resolution of the gel systems) to the sizes deduced from the DNA sequence of the r1 ORF, we are confident that a mature size protein is produced. Furthermore, this protein is relatively abundant in E. coli, indicating that the translation of the artificial r1 ORF universal code equivalent constructed must be efficient enough in E. coli despite major differences in codon usage bias.

The E. coli Synthesized r1 ORF Translational Product Shows a Specific Enzymatic Activity that Corresponds Exactly to All Predicted Properties of the Expected Genuine Mitochondrial Protein

The double strand endonuclease activity observed in the E. coli cells can be unambiguously attributed to the translation product of the omega⁺ r1 ORF since it follows the specificity requirements predicted for the intron transposase in mitochondria: it cleaves specifically the normal omega⁻ site but not the omega^n mutant sites; and the replacement of the omega⁺ r1 ORF by the omega^d r1 ORF abolishes the activity on the omega⁻ site. It follows that, in mitochondria, the r1 ORF product must, itself, be the active double strand endonuclease that generates a

double strand cut at the omega⁻ site to determine the insertion of the r1 intron at that site. No additional protein(s) or cofactor of either mitochondrial or nuclear origin is required in the specificity of this process.

It is notable that, in all our in vivo assays, the linearization of the plasmid population never goes to completion. This can be interpreted in several ways. First, the very construction of plasmids that simultaneously contain the r1 ORF and the target site for its product is likely to generate an equilibrium between the concentration of the enzyme and the number of linearized plasmids. The fact that only supercoiled plasmids can replicate probably maintains this equilibrium and explains our observation that the recombinant plasmids that express the active r1 ORF endonuclease are not rapidly lost in growing E. coli cultures on nonselective medium. Second, the E. coli cells may efficiently repair the linearized plasmids. However, the strain used was recA⁻ and the use of a lig⁻ mutant did not reveal significant differences with the lig⁺ strain. Third, the E. coli made protein may not be as active as the native mitochondrial protein due to the presence of two leucines instead of two threonines (at positions 123 and 156). Sequence comparisons with Kluyveromyces thermotolerans indicate, however, that these two threonines are not conserved, hence their probable unimportance. Last, there remains the interesting possibility that the double strand break observed represents an intermediate in the reaction catalyzed by the protein instead of a final product. This possibility is worth considering for it is remarkable that no degradation by unspecific E. coli exonucleases seems to occur at the break point of the linearized plasmids in our in vivo assays. This may indicate that the r1 ORF endonuclease itself protects the split omega⁻ site. Further in vitro tests are necessary to solve this question, but the limited proportion of linearized plasmids in E. coli parallels the very low fraction of mtDNA molecules having a double strand cut at the omega⁻ site in zygotes (Dujon et al., 1985; Zinn and Butow, 1985), again suggesting that no other proteins or cofactors are involved in mitochondria.

Specificity of the Double Strand Endonuclease Recognition Site

The overproduction of a double strand endonuclease in a cell should be deleterious to this cell if a recognition site (or sites) were present, by chance, in the chromosomes. Such deleterious effects have already been reported after overproduction of type II restriction endonucleases in yeast (Barnes and Rine, 1985) or in E. coli (Walder et al., 1984). In our experiments, however, no major effect on the growth rate of the E. coli cells containing the plasmid with the omega⁺ r1 ORF could be detected. One possible interpretation is that the recognition site for the r1 ORF product is large enough that very few or even no such site occurs at random in the E. coli chromosome (a sequence of 11 base pairs should occur once in the E. coli chromosome on a random basis). A relatively large size for the recognition site is also directly indicated in our experiments by the fact that an 8 nucleotide long sequence identical to the central part of the omega⁻ site (GGATAACA)

happens to be present in the *tac1* promoter of our expression plasmids. Yet, this site remains uncut when the *omega⁻* site is cut. The recognition site for the r1 ORF endonuclease must, therefore, extend beyond these 8 nucleotides forming the center of the *omega⁻* site. It may be worth noting that the recognition site at the MAT locus for the nuclear HO endonuclease, which bears homology to the *omega⁻* site, extends to approximately 15–16 bp (Kostriken and Heffron, 1984).

Possible Role of the r1 Intron Encoded Double Strand Endonuclease in Mitochondria

Our results show for the first time the enzymatic activity of an intron encoded protein. The specificity of action of our E. coli made protein directly demonstrates that the yeast mitochondrial r1 intron encodes a protein whose function is to generate a double strand cut at a specific site within the 21S rRNA gene. This triggers the insertion of that intron at that site, hence the self spreading of the intron. Whether the r1 ORF product is alone responsible for all steps of the duplicative insertion of the r1 intron cannot be yet decided. One interesting possibility may be that the class II intron ORF products recently shown as having characteristic features of reverse transcriptases (Michel and Lang, 1985) participate in the duplicative step after the double strand cut is generated at the *omega⁻* site. But the question remains of the actual advantage of synthesizing a very specific endonuclease when the recognition site for that endonuclease is not present in the *omega⁺* mitochondria (since it is split by the intron itself), whereas in the *omega⁻* mitochondria the site is present but not the protein. One possibility would be that the r1 ORF protein is synthesized in mitochondria only when the recognition site is present. Presence of this site may result from mating to an *omega⁻* cell (the origin of the phenomenon of "polarity of recombination" that lead to the discovery of this activity) or from the spontaneous loss of the r1 intron. Exact intron deletions have already been reported when selecting revertants of *mit⁻* mutants within the reading frames of some introns of the cob-box gene (Gargouri et al., 1983). If a similar excision of the r1 intron occurs spontaneously with a given frequency, then the r1 ORF activity would immediately cleave the DNA at the newly created *omega⁻* site and the intron would be reintroduced at that site from the other copies of mtDNA in the same cell. But another possibility is that the r1 ORF protein is preserved to create new genomic arrangements by introducing an intron at new locations. Although no such phenomenon has been observed in S. cerevisiae laboratory strains, the presence in the *oli1* gene of Kluyveromyces fragilis of an intron very closely resembling the r1 intron (Jacquier and Dujon, unpublished) strongly suggests such a possibility.

Experimental Procedures

Strains, Plasmids, and Growth Media
E. coli strains JM101 (Δlac pro thi supE F' traD36 proA⁺B⁺ lacIQ lacZ ΔM15) and JM103 (Δlac pro thi supE strA endA sbcB15 F' traD36 proA⁺B⁺ lacIQ lacZΔM15) were obtained from New England Biolabs; strain DH1 (F⁻ recA1 endA1 gyrA96 thi1 hsdR17(r⁻m⁺) supE44) origi-

nated from Dr. D. Hanahan; strain CSR603 (thr1 leuB1 proA2 argE3 thi1 phr1 λ⁻ ara14 lacY galK2 xyl5 man1 rps131 T633 supE44 recA1 uvrA6) was obtained from Dr. H. DeReuse.

The E. coli expression vector pDR540 was purchased from P. L. Biochemicals. M13mp8 and M13mp9 vectors were purchased from New England Biolabs. All other plasmids have been constructed in this work.

Standard E. coli growth media (LB, dYT, M9) were made according to Miller (1972). Ampicillin was added to a final concentration of 25 μg/ml when necessary.

Purification of Plasmid and M13 Phage DNAs
Large scale preparations of supercoiled plasmid DNA and minipreparations of plasmid DNA were made according to Maniatis et al. (1982). Replicative form and single-stranded form of M13 recombinants DNA were purified according to Messing (1983).

Synthesis and Purification of Oligonucleotides
Oligonucleotides were synthesized according to the solid phase phosphotriester method (Gait, 1980; Efimov et al., 1983) using a "SAM one" automated DNA synthesizer (Biosearch, USA). Deprotected oligonucleotides were purified by electrophoresis in 20% polyacrylamide denaturing gels (Maxam and Gilbert, 1980).

Determination of Dissociation Temperatures of the Synthetic Oligonucleotides
For each oligonucleotide, the exact dissociation temperature was first determined experimentally against the wild-type sequence in order to use optimal detection conditions of the mutants created at each mutagenesis cycle. Approximately 0.1 to 0.5 μg of purified single strand DNA from the appropriate M13 recombinants were put on 1 cm × 1 cm squares of nitrocellulose membrane (Millipore HAHY), air dried, and heated at 80°C under vacuum for 1–2 hr. Prehybridization and hybridizations were carried out at 10°C in 6× NET (1× NET is 150 mM NaCl, 15 mM Tris Cl, 1 mM EDTA at pH 7.5), 0.5% Nonidet P40 (FlukaAG, Switzerland), 1 mg/ml bovine serum albumin, 1 mg/ml Ficoll, 1 mg/ml polyvinylpyrrolidone, and 100 μg/ml of heat denatured sonicated herring sperm DNA. Hybridizations were carried out at 10°C for 4 to 12 hr using ³²P 5'-end-labeled oligonucleotides (approximately 10–50 ng of oligonucleotides with a specific radioactivity of 100–500 × 10⁶ cpm/μg were used for each experiment). After hybridization, three successive rinses were carried out at 0°C in 6× NET and 0.5% Nonidet P40 for approximately 15 min each. Each nitrocellulose square was then placed into a screw cap tube containing 15 ml of 6× NET, 0.5% Na dodecyl sulfate and put into a temperature controlled water bath. Temperature was raised by steps of 3°C every 5 min between predetermined limits and, at each step, a single square was removed and autoradiographed. The experimental dissociation temperature is defined as the first temperature step at which melting of the hybrid occurs (see Table 2). Comparison of these figures with the dissociation temperatures calculated from empirical formulae generally shows good agreement except for some oligonucleotides for which the actual thermal stability of the hybrids differs by 5–10°C from the calculated values. Dissociation temperatures were also determined against the mutant sequence (after the mutants were obtained) in order to estimate experimentally the increase in thermal stability due to the elimination of the different types of mismatches and to compare, again, the experimental figures to the calculated ones (Table 2). In all cases the temperature shifts are large enough (approximately 10°C to 20°C) to ensure proper detection of the mutation.

Oligonucleotide-Directed Multiple Mutagenesis
Each cycle of mutagenesis included the following steps. Step 1: approximately 0.1 μg (0.05 pmol) of purified single strand DNA from the appropriate M13 recombinant was annealed with 5 ng (1 pmol) of the universal primer (New England Biolabs pentadecamer) and 10–30 ng (2 pmol) of one (or two) D oligonucleotide previously phosphorylated at its 5' end using γ-³²P-ATP. Annealing was conducted in 20 μl of 10 mM Tris Cl, 50 mM NaCl, 20 mM MgCl₂, and 1 mM dithiothreitol at pH 7.8 by slowly decreasing the temperature from 85°C to room temperature. Step 2: elongation and ligation were conducted in 100 μl of the same buffer containing, in addition, 0.3 mM ATP, 0.3 mM of each dNTP, 1 mM spermidine, and 100 μg/ml bovine serum albumin. After addition

of 5 units of Klenow polymerase (Boehringer Mannheim) and 5–8 units of T4 DNA ligase (New England Biolabs) incubation was carried out at 15°C for 1 hr and then at 10°C overnight. The efficiency of the elongation and ligation were generally tested by gel electrophoresis of a restriction digest of a 5 μl aliquot using either Hind III (mSCM101 series) or Hind III + Sal I (mSCM109 and mSCM110 series) followed by autoradiography (notice that only the 5' end of the D oligonucleotide is ^{32}P-labeled such that the presence of a radioactive band corresponding to the insert size indicates an appropriate elongation and ligation). Step 3: the complete elongation–ligation mixtures were then added with 200 μl of TCM buffer (10 mM Tris Cl, 10 mM CaCl$_2$, 10 mM MgCl$_2$ at pH 7.8) and used to transform CaCl$_2$ treated E. coli cells (both JM101 and JM103 strains have been used). After transformation, the cells were incubated for 2 hr at 37°C in 1 ml LB medium prior to plating. This allows the completion of at least one replication cycle of M13 phages within the infected cells and largely eliminates the formation of mixed plaques. Serial dilutions were then plated on LB plates using 1.2% melted bacto agar, 0.1 ml of fresh E. coli cells, 800 μg/plate X-gal, and 400 μg/plate IPTG, and the plates were incubated at 37°C for approximately 12 hr.

Screening of mutants was done by plaque hybridization. Plates containing approximately 100–1000 plaques were covered with a nitrocellulose filter and phage diffusion was allowed for a few minutes. Filters were then treated with 100 mM NaOH, 1.5 M NaCl for 30 sec, and with 200 mM Tris Cl, 2× SSC at pH 7.5 for 1 min. Filters were baked and hybridized using the same ^{32}P-5'-end-labeled D oligonucleotide used for the mutagenesis. Hybridization was carried out in the same conditions as for the determination of the dissociation temperature except that the last rinse was made at 3°C above the dissociation temperature previously determined for the wild-type sequence. Four to eight positive plaques were taken at random from each mutagenesis cycle, inoculated into JM101 or JM103 cells, and the sequence of the mutagenized segment was determined. One mutant was then chosen and the sequence of its entire insert determined prior to using this mutant for the next mutagenesis cycle to ensure that no unexpected mutation occurred (this was facilitated by the use of the various D oligonucleotides as specific primers to ensure continuous reading within the entire insert). Notice that contrary to a previous report using a similar methodology (Osinga et al., 1983), we did not observe such unexpected mutations in our experiments. One possibility for this difference could be that, in our case, the elongation reaction was carried out at a lower temperature.

DNA Sequence Determinations
Single strand DNA from M13 recombinants was used for chain termination sequencing (Sanger et al., 1977).

Maxicells
After transformation of the E. coli strain CSR603 with the appropriate recombinant plasmid, a single ampicillin resistant transformant was picked up. All subsequent steps for protein labeling in the maxicells were as previously described (Sancar et al., 1979).

In Vivo Endonuclease Assays
Assay in mid-log cultures is as follows: 0.1 ml of fresh overnight cultures of E. coli cells containing the appropriate plasmid were inoculated into 10 ml LB medium containing ampicillin and incubated at 37°C for 2 hr. Cultures were then separated into two 5 ml aliquots, one added to IPTG to a final concentration of 10 mM, and both aliquots were further incubated for 2 hr. DNA was then extracted in parallel from each aliquot using the gentle procedure below.

Gel transfer hybridizations were carried out according to Southern (1975). Prehybridization and hybridization were performed at 65°C in 6× SSC (1× SSC is 150 mM NaCl, 15 mM trisodium citrate), 0.1% NaDodecyl sulfate, 1 mg/ml bovine serum albumin, 1 mg/ml ficoll, 1 mg/ml polyvinylpyrrolidone, and 100 μg/ml of heat denatured sonicated herring sperm DNA, using ^{32}P nick-translated mtDNA of strain HC9-7/J262 as a probe. After hybridization, nitrocellulose filters were rinsed twice in 2× SSC at room temperature, twice in 6× SSC, 0.1% NaDodecyl sulfate at 65°C, and finally in 1× SSC at room temperature.

The gentle preparation of total DNA used to purify supercoiled and linearized plasmids simultaneously is as follows: cells of appropriate cultures are centrifuged and rinsed once in 50 mM Tris Cl, 25% su-

crose (pH 8.0). Cells are then resuspended in 0.15 ml of the same buffer containing 2.5 mg/ml of lysozyme (Sigma) and kept on ice for 10 min, followed by the addition of 0.05 ml of 500 mM EDTA and a further incubation for 5 min on ice. Finally, 0.25 ml of triton solution (0.1% Triton X100, 62 mM EDTA, 50 mM Tris Cl, pH 8.0) is added, followed by a 10 min incubation on ice and a 10 min centrifugation at 15,000 rpm. The supernatant is carefully removed, placed in a new Eppendorf tube and added with a half volume of 7.5 M ammonium acetate and three volumes of ethanol, chilled at −70°C and centrifuged. The pellet is redissolved once in 0.2 ml of 300 mM sodium acetate, ethanol precipitated to eliminate ammonium ions, and finally rinsed with ethanol and vacuum dried. The final pellet is redissolved in 0.1 ml of TE buffer.

References

Barnes, G., and Rine, J. (1985). Regulated expression of endonuclease EcoRI in Saccharomyces cerevisiae: nuclear entry and biological consequences. Proc. Natl. Acad. Sci. USA 82, 1354–1358.

Barrell, B. G., Bankier, A. T., and Drouin, J. (1979). A different genetic code in human mitochondria. Nature 282, 189–194.

Bonitz, S. G., Berlani, R., Corruzi, G., Li, M., Macino, G., Nobrega, F., Nobrega, M., Thalenfeld, B. E., and Tzagoloff, A. (1980). Codon recognition rules in yeast mitochondria. Proc. Natl. Acad. Sci. USA 77, 3167–3170.

Burke, J. M., and RajBhandary, U. L. (1982). Intron within the large rRNA gene of N. crassa: a long open reading frame and a consensus sequence possibly important in splicing. Cell 31, 509–520.

Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E., and Slonimski, P. P. (1983). An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of the cytochrome oxidase in S. cerevisiae. Cell 35, 733–742.

Church, G., and Gilbert, W. (1980). Yeast mitochondrial intron products required in trans for RNA splicing. In Mobilization and Reassembly of Genetic Information, D. R. Joseph, J. Schultz, W. A. Scott, and R. Werner, eds. (New York: Academic Press), pp. 379–395.

Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzochio, C. (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. Nature 300, 719–724.

Dujon, B. (1980). Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles as the ω and rib-1 loci. Cell 20, 185–197.

Dujon, B. (1981). Mitochondrial genetics and functions. In Molecular Biology of the Yeast, J. N. Strathern, E. W. Jones, and J. R. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 505–635.

Dujon, B. (1983). Mitochondrial genes, mutants, and maps: a review. In Mitochondria 1983, R. J. Schweyen, K. Wolf, and F. Kaudewitz, eds. (Berlin: W. de Gruyter and Co.), pp. 1–24.

Dujon, B., and Jacquier, A. (1983). Organization of the mitochondrial 21S rRNA gene in Saccharomyces cerevisiae: mutants of the peptidyl

transferase centre and nature of the omega locus. In Mitochondria 1983, R. J. Schweyen, R. J. Wolf, and F. Kaudowitz, eds. (Berlin: W. de Gruyter and Co.), pp. 389–403.

Dujon B., Bolotin-Fukuhara, M., Coen, D., Deutsch, J., Netter, P., Slonimski, P. P., and Weill, L. (1976). Mitochondrial genetics. XI. Mutations at the mitochondrial locus omega affecting the recombination of mitochondrial genes in Saccharomyces cerevisiae. Mol. Gen. Genet. *143*, 131–165.

Dujon, B., Cottarel, G., Colleaux, L., Betermier, M., Jacquier, A., d'Auriol, L., and Galibert, F. (1985). Mechanism of integration of an intron within a mitochondrial gene: a double strand break and the transposase function of an intron encoded protein as revealed by in vivo and in vitro assays. In Achievements and Perspective in Mitochondrial Research, F. Palmieri, ed. (Amsterdam: Elsevier), in press.

Efimov, V. A., Burykova, A. A., Reverdatto, S. V., Chakhmakhcheva, O. G., and Ovchinnikov, Y. A. (1983). Rapid synthesis of long-chain deoxyribonucleotides by the N-methylimidazolide phosphotriester method. Nucl. Acids Res. *11*, 8369–8387.

Fox, T. D. (1979). Five TGA stop codons occur within the translated sequence of the yeast mitochondrial gene for cytochrome oxidase subunit II. Proc. Natl. Acad. Sci. USA *76*, 6534–6538.

Gait, M. J. (1980). Polymer-Supported Reactions in Organic Synthesis, P. Hodge, and D. C. Sherrington, eds. (New York: John Wiley), p. 435.

Gargouri, A., Lazowska, J., and Slonimski, P. P. (1983). DNA splicing of introns: a general way of reverting intron mutations. In Mitochondria 1983, R. J. Schweyen, K. Wolf, and F. Kaudewitz, eds. (Berlin: W. de Gruyter and Co.), pp. 259–268.

Gouy, M., and Gautier, C. (1982). Codon usage in bacteria: correlation with gene expressivity. Nucl. Acids Res. *10*, 7055–7074.

Grivell, L. (1983). Mitochondrial gene expression 83. In Mitochondria 1983, R. J. Schweyen, K. Wolf, and F. Kaudewitz, eds. (Berlin: W. deGruyter and Co.), pp. 25–45.

Guiso, N., Dreyfus, M., Siffert, O., Danchin, A., Spirydakis, A., Gargouri, A., Claisse, M., and Slonimski, P. P. (1984). Antibodies against synthetic oligopeptides allow identification of the mRNA maturase encoded by the second intron of the yeast cob-box gene. EMBO J. *3*, 1769–1772.

Hudspeth, M. E. S., Ainlay, W. M., Shurnard, D. S., Butow, R. A., and Grossman, L. I. (1982). Location and structure of the *Var1* gene on yeast mitochondrial DNA: nucleotide sequence of the *40.0* allele. Cell *30*, 617–626.

Jacq, C., Lazowska, J., and Slonimski, P. P. (1980). Sur un nouveau mécanisme de la régulation de l'expression génétique. Comp. Rend. Acad. Sci. Paris *290*, 89–92.

Jacq, C., Banroques, J., Becam, A. M., Slonimski, P. P., Guiso, N., and Danchin, A. (1984). Antibodies against a fused "lac Z yeast mitochondrial intron" gene product allow identification of the mRNA maturase encoded by the fourth intron of the yeast cob-box gene. EMBO J. *3*, 1567–1572.

Jacquier, A., and Dujon, B. (1983). The intron of the mitochondrial 21S rRNA gene: distribution in different yeast species and sequence comparison between Kluyveromyces thermotolerans and Saccharomyces cerevisiae. Mol. Gen. Genet. *192*, 487–499.

Jacquier, A., and Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell *41*, 383–394.

Kostriken, R. G., and Heffron, F. (1984). The product of the HO gene is a nuclease: purification and characterization of the enzyme. Cold Spring Harbor Symp. Quant. Biol. *49*, 89–96.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147–157.

Lathe, R. (1985). Synthetic oligonucleotide probes deduced from amino acid sequence data. J. Mol. Biol. *183*, 1–12.

Lazowska, J., Jacq, C., and Slonimski, P. P. (1980). Sequence of introns and flanking exons in wild-type and *box3* mutants of cytochrome b reveals an interlaced splicing protein coded by an intron. Cell *22*, 333–348.

Li, M., and Tzagoloff, A. (1979). Assembly of the mitochondrial membrane system: sequences of yeast mitochondrial valine and an unusual threonine tRNA gene. Cell *18*, 47–53.

Macino, G., Corruzi, G., Nobrega, F., Li, M., and Tzagoloff, A. (1979). Use of the UGA terminator as a tryptophan codon in yeast mitochondria. Proc. Natl. Acad. Sci. USA, *76*, 3784–3785.

Macreadie, I. G., Scott, R. M., Zinn, A. R., and Butow, R. A. (1985). Transposition of an intron in yeast mitochondria requires a protein encoded by that intron. Cell *41*, 395–402.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual, (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base specific cleavages. In Methods in Enzymology, R. Wu, L. Grossman, and K. Moldave, eds. (New York: Academic Press), *65*, pp. 499–560.

Messing, J. (1983). New M13 vectors for cloning. In Methods in Enzymology *101*, R. Wu, L. Grossman, and K. Moldave, eds. (New York: Academic Press), pp. 20–78.

Michel, F., and Dujon, B. (1983). Conservation of RNA secondary structure in two intron families including mitochondrial, chloroplast and nuclear encoded members. EMBO J. *2*, 33–37.

Michel, F., and Lang, B. F. (1985). Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. Nature *316*, 641–643.

Michel, F., Jacquier, A., and Dujon, B. (1982). Comparison of fungal mitochondrial introns reveals extensive homology in RNA secondary structure. Biochimie *64*, 867–881.

Miller, J. H. (1972). Experiments in Molecular Biology. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Osinga, K. A., van der Blieck, A. M., van der Horst, G., Groot Koerkamp, M. J. A., and Tabak, H. F. (1983). In vitro site directed mutagenesis with synthetic oligonucleotides yields unexpected deletions insertions at high frequency. Nucl. Acids Res. *11*, 8595–8608.

Sancar, A., Hack, A. M., and Rupp, W. D. (1979). Simple method for identification of plasmid-coded protein. J. Bacteriol. *137*, 692–693.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Strathern, J. N., Klar, A. J. S., Hicks, J. B., Abraham, J. A., Ivy, J. M., Nasmith, K. A., and McGill, C. (1982). Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the *MAT* locus. Cell *31*, 183–192.

Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K., and Wallace, R. B. (1981). ICN–UCLA symposium. Mol. Cell. Biol. *23*, 682–693.

Tabak, H. F., van der Laan, J., Schouten, J. P., van Boom, J. H., and Veeneman, G. H. (1981). Use of synthetic DNA oligonucleotide to probe the precision of RNA splicing in yeast mitochondrial petite mutant. Nucl. Acids Res. *9*, 4475–4483.

van der Horst, G., and Tabak, H. F. (1985). Self-splicing of yeast mitochondrial ribosomal and messenger RNA precursors. Cell *40*, 759–766.

Walder, R. Y., Walder, J. A., and Donelson, J. E. (1984). The organisation and complete nucleotidic sequence of the Pst I restriction-modification system. J. Biol. Chem. *259*, 8015–8026.

Waring, R. B., and Davies, R. W. (1984). Assessment of a model for intron RNA secondary structure relevant to RNA self splicing: a review. Gene *28*, 277–291.

Weiffenbach, B., Rogers, D. T., Haber, J. E., Zoller, M., Russell, D. W., and Smith, M. (1983). Deletion and single base pair changes in the yeast mating type locus that prevent homothallic mating type conversion. Proc. Natl. Acad. Sci. USA *80*, 3401–3405.

Zinn, A. R., and Butow, R. A. (1985). Nonreciprocal exchange between alleles of the yeast mitochondrial 21S rRNA gene: kinetics and involvement of a double-strand break. Cell *40*, 887–895.